Role of boundary constraints in DNA cyclization

Alexei V. Tkachenko
Department of Physics and Michigan Center for Theoretical Physics, University of Michigan, 450 Church Str., Ann Arbor, 48109 MI, USA

We modify the classical Shimada-Yamakawa theory of DNA looping by generalizing the form of boundary constraints. This generalization is important in the context of DNA cyclization experiments since it mimics the reduced local rigidity of the "nicked" DNA loop. Our results indicate that the non-trivial boundary constraints may be responsible for the existing dramatic discrepancy between various DNA cyclization experiments. The developed effective Hamiltonian method may be extended to even broader class of DNA looping problems.

PACS numbers: 87.14.Gg, 87.15.La, 87.15.Aa

Loop formation in double-stranded DNA (dsDNA) is essential for such important biological processes as regulation of gene expression and DNA packaging into nucleosomes. In the first case, the loop is induced by the interactions between transcription factor proteins bound to different sites along the DNA chain [1]-[3]. In the case of nucleosome, DNA wraps around the near-cylindrical histone octamer. The classical theory of looping, based on the elastic description of DNA, was proposed more than two decades ago by Shimada and Yamakawa (SY) [4]. There are however multiple indications that the original theory is not sufficiently adequate for describing the real experimental situation.

Partially, the deviations can be attributed to the complexity of the actual *in vivo* problem. However, a large discrepancy is also reported in recent *in vitro* experiments on DNA cyclization [5][6]. In these experiment, looping is induced by hybridization of mutually complementary ssDNA ends of the chain (see Figure 1). Even in this relatively simple case, the agreement between theory and experiment is a highly controversial issue.

On the one hand, older cyclization experiments by Shore et al do agree with SY model [7][8]. In fact, those data were used in the original SY paper to support their model. On the other hand, Cloutier and Widom reported that the discrepancy between experimental looping probability and SY results may reach two to three orders of magnitude [5]. Their paper was followed by the work of Vologodskii lab, in which the agreement between theory and experiment was confirmed [9]. More recently, Cloutier and Widom reiterated their claim, and also reported that the twist-related oscillations of the looping probability are strongly reduced compared to the SY prediction [6]. This controversy inspired a new interest to the problem among theorists. In particular, it was suggested that strong bending may induce local structural defects such as "kinks" or ssDNA "bubbles" (which act as "soft" kinks)[10]-[12]. At present, there is no direct evidence to support any of these models, and they do not resolve the conflict between the different experiments.

In this communication, we propose an alternative explanation to the existing controversy. We argue that the DNA cyclization involve more complicated boundary (i.e. terminal) constraints then it is traditionally believed. In order to solve the problem with the modified boundary conditions, we develop an effective Hamiltonian method. The approach is potentially applicable beyond the scope of this work, e.g. for study of protein-mediated looping [13].

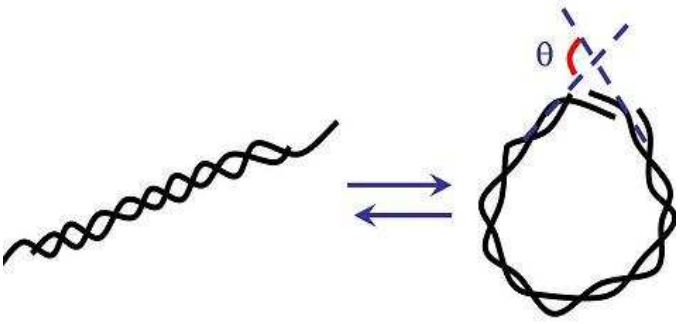

FIG. 1: Scheme of DNA cyclization experiment. Looping is caused by hibridization of mutually complementary ssDNA ends.

The SY theory of looping starts with dsDNA modelled as an elastic rod subjecxted to thermal fluctuations, which is essensially equivalient to Worm Like Chain (WLC) model [14]-[15]:

$$H_0 = k_B T \int_0^L ds \left[\frac{l_p}{2} \left(\frac{\partial \hat{\mathbf{t}}}{\partial s} \right)^2 + \frac{l_t}{2} \left(\frac{\partial \psi}{\partial s} \right)^2 \right] \qquad (1)$$

Here, the first term describes the bending elasticity with modulus proportional to dsDNA persistence length $l_p \approx 50nm$, the second term represents torsional elasticity, with the modulus proportional to twist persistence length, $l_t \simeq 100nm$. $\hat{\mathbf{t}}$ is unit tangent vector of the chain, and ψ is twist angle, both functions of position along the chain, s.

From the experimental and biological points of view, the most important characteristics of the looping problem is the so called J-factor. It has a meaning of an effective concentration of one end of the loop in the vicinity of the other, at the open configuration. J-factor can be related to equilibrium probability for the loop to be closed,

$P^{(loop)}$, as well as to the closing/opening times (τ_{cl}, τ_{op}):

$$\frac{P^{(loop)}}{1-P^{(loop)}} = \frac{\tau_{op}}{\tau_{cl}} = \frac{J}{c_0}\exp\left(\frac{\varepsilon}{k_BT}\right) \qquad (2)$$

Here ε is the mutual affinity of the loop terminals, and $c_0 = 1M$ is the standard reference concentration. Affinity ε can be independently determined from the experiment on dimerization of free reacting terminal groups in solution.

Shimada and Yamakawa calculated J-factors for two important cases: circular loops with completely aligned ends ($\hat{\mathbf{t}}(L) = \hat{\mathbf{t}}(0)$):

$$J_0 \approx \frac{32\pi^3}{l_P^3}\left(\frac{l_P}{L}\right)^6 \exp\left(-\frac{2\pi^2 l_p}{L} - \frac{L}{4l_P}\right), \qquad (3)$$

and loops with unconstrained orientations of the end segments:

$$J_{free} \approx \frac{110}{l_P^3}\left(\frac{l_P}{L}\right)^5 \exp\left(-\frac{14.04 l_p}{L} - \frac{L}{4l_P}\right). \qquad (4)$$

These results are obtained in the limit of short loops. However, their range of applicability extends up to $L \sim 10l_P$ (i.e. 1500 bp). The effects of torsional constraints which are not included in Eqs.(3)-(4), result in an additional factor with a pronounced oscillatory behavior with the period of DNA helical turn, 10.5 bp.

DNA loop formed in a reversible cyclization experiments (i.e. before ligation) is not identical to a circular DNA. It is "nicked" in two points corresponding to the ends of the DNA strands (see Figure 1). These singular points are expected to have a greater flexibility then the rest of the chain [16]. The effective local flexibility must strongly depend on the base-stacking interactions. We will describe the coupling between orientations of the end segments of the loop with the following minimal model:

$$\frac{H_{end}}{kT} = \frac{\kappa\theta^2}{2} + \frac{\kappa'(\Delta\psi)^2}{2} \qquad (5)$$

Here θ is the angle between directions of tangent vectors, $\hat{\mathbf{t}}(0)$ and $\hat{\mathbf{t}}(L)$, and $\Delta\psi$ is the relative twist of the two segments. The two parameters, κ and κ' should be of the same order for the given sequence. Since the characteristic stacking energy is of order of kT, we expect $\kappa \simeq \kappa' \sim 1$.

In order to calculate $J-$ factor with this modified boundary constraints, we develop an *effective Hamiltonian method* (EHM) for the problem. Effective Hamiltonian can be introduced as a free energy of the loop parameterized with orientations of its end segments. If we neglect any torsional constraints (i.e. assume $\kappa' = 0$), the Hamiltonian of a closed DNA loop can be written as a function of azimuthal and polar angles, θ and φ of tangent vector $\hat{\mathbf{t}}(L)$, with respect to the vector triad at the other end of the loop, $s = 0$:

$$H = H_{loop}(\theta,\varphi) + \frac{\kappa\theta^2}{2} \qquad (6)$$

In order to construct $H_{loop}(\theta,\varphi)$, we first discuss the ground state energy of the loop. Since the original Hamiltonian of WLC model is formally equivalent to Lagrangian of symmetrical spinning top, finding its ground state is an integrable mechanical model. However, finding the exact solution for a particular set of boundary conditions, requires an inversion of incomplete Elliptical functions, and therefore it is not practical [13]. Instead, we consider the problem in the vicinity of circular loop configuration (which corresponds to $\hat{\mathbf{t}}(0) = \hat{\mathbf{t}}(L)$). In this limit, the ground state energy of a 2D (planar) loop can be written as an expansion in θ:

$$\frac{E_{loop}(\theta)}{kT} \approx \frac{l_P}{L}\left(2\pi^2 + \beta\theta + \frac{\gamma\theta^2}{2}\right) + O\left(\theta^3\right), \qquad (7)$$

where the exact values of the coefficients are: $\beta = 2\pi$; $\gamma = 3$. Already this expression gives an excellent description of the global behavior of function $E_{loop}(\theta)$, not limited to the near vicinity of circular loop ($\theta = 0$). In particular, it predicts the minimal energy, $E_{\min}/kT = (4\pi^2/3)l_p/L \approx 13.16 l_p/L$, which is very close to the results of complete numerical solution of the problem: $E_{\min}/kT \approx 14.04 l_p/L$. Nevertheless, due to the strong exponential dependence of J-factor on the elastic energy, we need a further improvement of the analytic formula for $E_{loop}(\theta)$. We achieve this by treating the coefficient γ in as a free parameter, and adjusting it to value $\gamma = 3.46$. This leads to an exact matching of the minimum energy. The overall behavior of approximate expression, Eq.(7), becomes nearly indistinguishable from the exact result (see Figure 2).

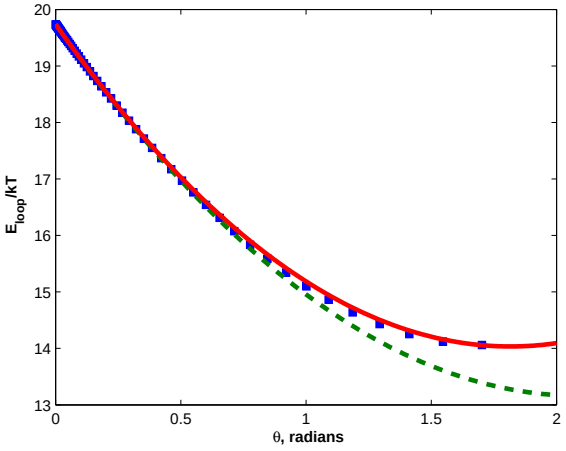

FIG. 2: Comparison of the Effective Hamiltonian (solid line), to the exact result for the elastic energy of a loop (squares). Dashed line shows the extrapolation from point $\theta = 0$, without any adgustment to model parameter γ (i.e $\gamma = 3$).

Since the change in the elastic energy is the dominant correction to the loop free energy, we can write the effective Hamiltonian of the loop in the following form:

$$\frac{H_{loop}(\theta,\varphi)}{kT} \approx \frac{F_0}{kT} + \frac{l_P}{L}\left(2\pi\theta\cos\varphi + \frac{\left(\gamma+\gamma'\sin^2\varphi\right)\theta^2}{2}\right), \tag{8}$$

Here, the earlier expression, Eq. (7), has been generalized for the case of 3D loops, and additional fine-tuning parameter, γ' has been introduced. F_0 is free energy of the circular loop.

The combined effective Hamiltonian Eq. (6), reaches minimum at $\theta = \theta^* = -2\pi/(\gamma + \kappa L/l_p)$, and $\varphi = 0$. We now expand it in the vicinity of that point:

$$\frac{H(\theta,\varphi)}{kT} \approx \frac{F_0}{kT} + \frac{2\pi^2 l_P}{L(\gamma+\kappa L/l_p)} + \left(\frac{\gamma l_P}{L}+\kappa\right)\frac{(\theta-\theta^*)^2}{2} + \tag{9}$$

$$+\frac{4\pi^2 l_P}{L(\gamma+\kappa L/l_p)}\left[1+\frac{\gamma'}{(\gamma+\kappa L/l_p)}\right]\frac{\varphi^2}{2}$$

This quadratic expansion allows one to calculate the overall free energy of the looped state in Gaussian approximation, and hence obtain J-factor for arbitrary coupling κ :

$$J_\kappa = J_0 \frac{\int \mathrm{e}^{-(H-F_0)/kT} d\Omega}{\int \mathrm{e}^{-H_{end}/kT} d\Omega} \tag{10}$$

The result of this calculation can be well approximated by the following analytic expression (the torsional effects are omitted):

$$J_\kappa \approx J_0(L/l_p)\frac{(2\kappa+1)L}{4\pi l_p}\sqrt{1+\frac{1}{1+2\kappa L/\gamma l_p}}\times \tag{11}$$

$$\times \sin\left(\frac{2\pi}{\gamma+\kappa L/l_p}\right)\exp\left(\frac{2\pi^2}{(\gamma+\kappa L/l_p)}\frac{l_P}{L}\right)$$

This form generalizes of the original SY result The parameter γ' was tuned to the value $\gamma' = -\gamma/2$, to achieve a nearly exact matching with both SY limits, Eqs. (3) and (4).

To complete our calculation we now need to include effects of the torsional constrain. For short loops, the corresponding effective Hamiltonian is the sum of the elastic energy of the twisted chain and the torsional part of H_{end}:

$$H_{tors}(\Delta\psi) \approx \frac{l_t}{2L}\left[\Delta\psi + 2\pi\left(N-\frac{L}{h}\right)\right]^2 + \frac{\kappa'(\Delta\psi)^2}{2} \tag{12}$$

Here N is an integer *linking number,* and h is helix repeat of dsDNA. This leads to an additional torsion-related factor in the final result:

$$J = \frac{J_\kappa(L/l_p)}{\sqrt{\kappa' L/l_t + 1}} \sum_{N=-\infty}^{+\infty} \exp\left[\frac{2\pi^2(N-L/h)^2}{L/l_t+\kappa'^{-1}}\right] \tag{13}$$

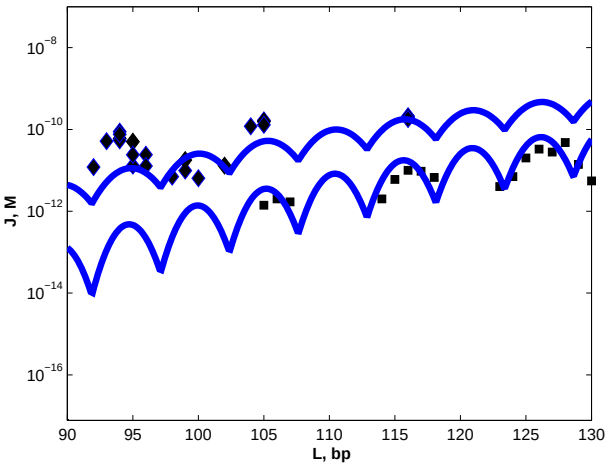

FIG. 3: Comparison of the theoretical result, Eq. (3) with experimental data from Refs. [5] (diamonds) and [9] (squares). The two curves correspond to weak ($\kappa = 1, \kappa' = 2$) and strong ($\kappa = \kappa' = 20$) orientational couplings, respectively.

The calculated J-factor is shown in Figure 3, as a function of the loop size. Remarkably, the two conflicting sets of experiments are both consistent with the model. In particular Cloutier-Widom and Vologoidskii lab data correspond to the regimes of weak ($\kappa = 1$, $\kappa' = 2$) and strong ($\kappa = \kappa' = 20$) terminal coupling, respectively. As expected, the two coupling parameters, κ and κ' are strongly correlated. This variation in value of κ is reasonable since the effective rigidity of a DNA "nick" is likely to have an exponential dependence on the local stacking energy. This strong dependence of the cyclization probability on the local sequence should not be confused with another sequence dependent effect associated with inhomogeneous intrinsic curvature and bending modulus of the chain [17]. In order to separate the two effects, the sequence of the mutually complementary terminal groups and that of the rest of the chain should be varied independently in the future experiments.

One can make several important conclusions based on our results. First, the traditional "circular loop" modelling of DNA cyclization is only justified for very strong coupling, $\kappa \gtrsim 10$. Second, in the regime of moderate coupling ($\kappa \simeq \kappa' \sim 1$) the effects of the two constraints are rather different. The overall shape of J-factor curve follows that of an orientationally unconstrained loop, Eq. (4), but the torsion-related oscillations, although greatly reduced, remain rather prominent. This is indeed consistent with the results of Ref. [6].

[1] R. Schleif, *Annu. Rev. Biochem.* **61**, 199 (1992).
[2] L. Finzi and J. Gelles, *Science* **267**, 378 (1995).
[3] F. Vanzi , C. Broggio , L. Sacconi , and F. S. Pavone, Nucl. Acids Res. **34**: 3409 (2006).
[4] J. Shimada, H. Yamakawa, *Macromolecules* **17**, 689 (1984).
[5] T. Cloutier, J. Widom, Molecular Cell **14**, 355 (2004)
[6] T. Cloutier and J. Widom, PNAS. **102**: 3645 (2005).
[7] D. Shore, J. Langowski, and R. L. Baldwin, *Proc. Natl. Acad. Sci. USA* **79**, 4833 (1981).
[8] D. Shore and R. L. Baldwin, *J Mol Biol.* **170,** 957 (1983)
[9] Q. Du, C. Smith, N. Shiffeldrim, M. Vologodskaia, and A. Vologodskii, PNAS **102**: 5397 (2005)
[10] J. Yan and J. F. Marko, Phys. Rev. Lett. 93, 108108 (2004)
[11] P. A. Wiggins, R. Phillips, and P. C. Nelson, Phys. Rev. E 71, 021909 (2005)
[12] P. Ranjith, P. B. Sunil Kumar, and Gautam I. Menon, Phys. Rev. Lett. 94, 138102 (2005)
[13] N. Douarche and S. Cocco, Phys. Rev. E **72**, 061902 (2005)
[14] O. Kratky and G. Porod, *Rec. Trav. Chim.* **68**, 1106 (1949).
[15] J. F. Marko, and E. Siggia, *Macromolecules* **28**, 8759 (1995).
[16] E. Protozanova, P. Yakovchuk and M. D. Frank-Kamenetskii, J. of Mol.Biology **342**, Issue 3, 775 (2004)
[17] Y.O. Popov, A.V. Tkachenko, cond-mat/0510302, to be published.